

Markov Equivalences for Subclasses of Loopless Mixed Graphs

KAYVAN SADEGHI

Department of Statistics, University of Oxford

ABSTRACT: In this paper we discuss four problems regarding Markov equivalences for subclasses of loopless mixed graphs. We classify these four problems as finding conditions for internal Markov equivalence, which is Markov equivalence within a subclass, for external Markov equivalence, which is Markov equivalence between subclasses, for representational Markov equivalence, which is the possibility of a graph from a subclass being Markov equivalent to a graph from another subclass, and finding algorithms to generate a graph from a certain subclass that is Markov equivalent to a given graph. We particularly focus on the class of maximal ancestral graphs and its subclasses, namely regression graphs, bidirected graphs, undirected graphs, and directed acyclic graphs, and present novel results for representational Markov equivalence and algorithms.

Key words: Bidirected graph, directed acyclic graph, m -separation, Markov equivalence, maximal ancestral graph, regression chain graph, summary graph, undirected graph.

1 Introduction

Introduction and motivation. In graphical Markov models several classes of graphs have been used in recent years. A common feature of all these graphs is that their nodes correspond to random variables, and they represent conditional independence statements of the node set by specific interpretations of missing edges.

These graphs contain up to three different types of edges. Sadeghi & Lauritzen (2011) gathered most classes of graphs defined in the literature under a unifying class of loopless mixed graphs (LMGs). These contain *Summary graphs* (SGs) (Wermuth, 2011), *(maximal) ancestral graphs* (MAGs) (Richardson & Spirtes, 2002), *acyclic directed mixed graphs* (ADMGs) (Spirtes *et al.*, 1997), regression chain graphs (RCGs) (Cox & Wermuth, 1993; Wermuth & Cox, 2004; Wermuth & Sadeghi, 2011), *undirected* or *concentration graphs* (UGs) (Darroch *et al.*, 1980; Lauritzen, 1996), *bidirected* or *co-variance graphs* (BGs) (Cox & Wermuth, 1993; Wermuth & Cox, 1998), and directed acyclic graphs (DAGs) (Kiiveri *et al.*, 1984; Lauritzen, 1996).

For the above graphs, in general, two graphs of different types or even two graphs of the same type may induce the same independencies. Such graphs are said to be *Markov equivalent*. It is important to explore the similar characteristics of Markov equivalent

graphs, and to find the ways of generating graphs of a certain type with the same independence structure from a given graph.

Some questions for Markov equivalences. There are four main questions regarding Markov equivalence for different types of graphs:

- 1) **Internal Markov equivalence:** The first natural question that arises in this context is regarding when two graphs of the same type are Markov equivalent. This question may be answered for DAGs, MAGs, or other subclasses of LMGs.
- 2) **External Markov equivalence:** In addition to Markov equivalence for graphs of the same type, one can discuss Markov equivalence between two graphs of different types.
- 3) **Representational Markov equivalence:** Before checking external Markov equivalence, however, it is essential to check whether and under what conditions a graph of a certain type can be Markov equivalent to a graph of another type.
- 4) **Algorithms:** One can also present some algorithms to generate a graph of a certain type that is Markov equivalent to a given graph of a different type.

In this paper we gather and simplify the existing results in the literature for internal and external Markov equivalences, and give novel results for representational Markov equivalence and algorithms.

Some earlier results on Markov equivalence for graphs. Results concerning Markov equivalence for different classes of graphs have been obtained independently in the statistical literature on specifying types of multivariate statistical models, and in the computer science literature on deciding on special properties of a given graph or on designing fast algorithms for transforming graphs. In the literature on graphical Markov models two of the early results concerning Markov equivalence for DAGs and chain graphs were respectively given in Verma & Pearl (1990) and Frydenberg (1990). Two of the later results by Zhao *et al.* (2004) and Ali *et al.* (2009) respectively provided theoretically neat and computationally fast conditions for Markov equivalence for maximal ancestral graphs.

Besides these, Pearl & Wermuth (1994) provided conditions for Markov equivalence for bidirected graphs and DAGs. Spirtes & Richardson (1997) gave some conditions for Markov equivalence for maximal ancestral graphs, in which the polynomial computational complexity claim was wrong.

Efficient algorithms for deciding whether a UG can be oriented into a DAG became available in the computer science literature under the name of perfect elimination orientations; see Tarjan & Yannakakis (1984), whose algorithm can be run in $O(|V| + |E|)$.

Another such linear algorithm can be found in Blair & Peyton (1992). An algorithm for generating a Markov equivalent DAG from a bidirected graph is the special case of the algorithm given in Zhao *et al.* (2004).

Structure of the paper. In the next section we define the unifying class of LMGs, and provide some basic graph theoretical definitions needed for our results.

In Section 3 we present the subclasses of LMGs, and we formally define the subclasses of interest in this paper. We also define a so-called separation criterion, called *m*-separation, to provide an interpretation of independencies for the graphs.

In Section 4 we formally define Markov equivalence, define maximality and explain its importance for Markov equivalences, and motivate why we consider Markov equivalence for the class of MAGs.

In Section 5 we gather the conditions existing in the literature for internal Markov equivalence for the class of MAGs and its subclasses, and give conditions for their external Markov equivalence.

In Section 6 we discuss the conditions for representational Markov equivalence for MAGs and its subclasses to a specific subclass, and we also provide algorithms to generate a Markov equivalent graph of a specific type to a given graph of another type when the conditions for representational Markov equivalence are satisfied. In each subsection we deal with different subclasses of MAGs: DAGs in Section 6.1, UGs and BGs in Section 6.2, and RCGs in Section 6.3.

In Section 7 we summarise the results, presented in the paper.

2 Loopless mixed graphs

Graphs. A *graph* G is a triple consisting of a *node* set or *vertex* set V , an *edge* set E , and a relation that with each edge associates two nodes (not necessarily distinct), called its *endpoints*. A *loop* is an edge with the same endpoints. When nodes i and j are the endpoints of an edge, these are *adjacent* and we write $i \sim j$. We say the edge is *between* its two endpoints. We usually refer to a graph as an ordered pair $G = (V, E)$. Graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are called *equal* if $(V_1, E_1) = (V_2, E_2)$. In this case we write $G_1 = G_2$.

Notice that the graphs that we use in this paper (and in general in the context of graphical models) are so-called *labeled graphs*, i.e. every node is considered a different object. Hence, for example, graph $i \text{ --- } j \text{ --- } k$ is not equal to $j \text{ --- } i \text{ --- } k$.

Definition of loopless mixed graph. Sadeghi & Lauritzen (2011) gathered most graphs in the literature of graphical models under the definition of *loopless mixed graph*, which is

a graph that contains three types of edges denoted by arrows, arcs (two-headed arrows), and lines (full lines) and does not contain any loops.

Basic definitions for LMGs. We say that i is a *neighbour* of j if these are endpoints of a line, and i is a *parent* of j if there is an arrow from i to j . We also define that i is a *spouse* of j if these are endpoints of an arc. We use the notations $\text{ne}(j)$, $\text{pa}(j)$, and $\text{sp}(j)$ for the set of all neighbours, parents, and spouses of j respectively. In the cases of $i \rightarrow j$ or $i \leftrightarrow j$ we say that there is an arrowhead pointing to (at) j .

A *subgraph* of a graph G_1 is graph G_2 such that $V(G_2) \subseteq V(G_1)$ and $E(G_2) \subseteq E(G_1)$ and the assignment of endpoints to edges in G_2 is the same as in G_1 . An *induced subgraph* by nodes $A \subseteq V$ is a subgraph that contains all and only nodes in A and all edges between two nodes in A . A subgraph induced by edges $A \subseteq E$ is a subgraph that contains all and only edges in A and all nodes that are endpoints of edges in A . We denote the subgraphs induced by arrows, arcs, and lines in a graph H by $H[\rightarrow]$, $H[\leftrightarrow]$, and $H[—]$ respectively.

A *walk* is a list $\langle v_0, e_1, v_1, \dots, e_k, v_k \rangle$ of nodes and edges such that for $1 \leq i \leq k$, the edge e_i has endpoints v_{i-1} and v_i . A *path* is a walk with no repeated node or edge. We denote a path by an ordered sequence of node sets. We say a path is *between* the first and the last nodes of the list in G . We call the first and the last nodes *endpoints* of the path and all other nodes *inner nodes*.

A *cycle* in a graph G is a simple subgraph whose nodes can be placed around a circle so that two nodes are adjacent if these appear consecutively along the circle.

A path (or cycle) is *direction preserving* if all its edges are arrows pointing to one direction. If a direction-preserving path is from a node j to a node i then j is an *ancestor* of i . We denote the set of ancestors of i by $\text{an}(i)$.

A graph is called *directed* if it only contains arrows. A directed graph is *acyclic* if it has no direction-preserving cycle.

A *chord* is an edge between two non-adjacent nodes on the cycle. A cycle is *chordless* if it has no chords. The notation C_n is used for a chordless cycle with n nodes. Notice that C_n can contain different types of edges, so it represents a class of graphs rather than a single graph. We call a graph *chordal* if it has no C_n , $n \geq 4$, as an induced subgraph. We also use the notation P_n for a *chordless* or *minimal* path with n nodes, i.e. a path that has no edge between two non-adjacent nodes on the path.

A *V-configuration* is a path with three nodes and two edges. Notice that originally and in most texts, e.g. in Kiiveri *et al.* (1984), the endpoints of a V-configuration is by definition not adjacent. In this paper we call such V-configurations *unshielded*.

In a mixed graph the inner node of three V-configurations $i \rightarrow t \leftarrow j$, $i \leftrightarrow t \leftarrow j$, and $i \leftrightarrow t \leftrightarrow j$ is a *collider* and the inner node of all other V-configurations is a

non-collider node on the V-configuration or more generally on a path of which the V-configuration is a subpath. We also call the V-configuration with collider or non-collider inner node a *collider* or *non-collider V-configuration* respectively. We may speak of a collider or non-collider node without mentioning the V-configuration or the path when this is apparent from the context. In the case of DAGs the only collider V-configuration $i \rightarrow t \leftarrow j$ is called a collision V-configuration.

3 Subclasses of loopless mixed graphs and their independence interpretation

Subclasses of LMGs The following diagram, presented in Sadeghi & Lauritzen (2011) illustrates the hierarchy regarding inclusions of subclasses of LMGs.

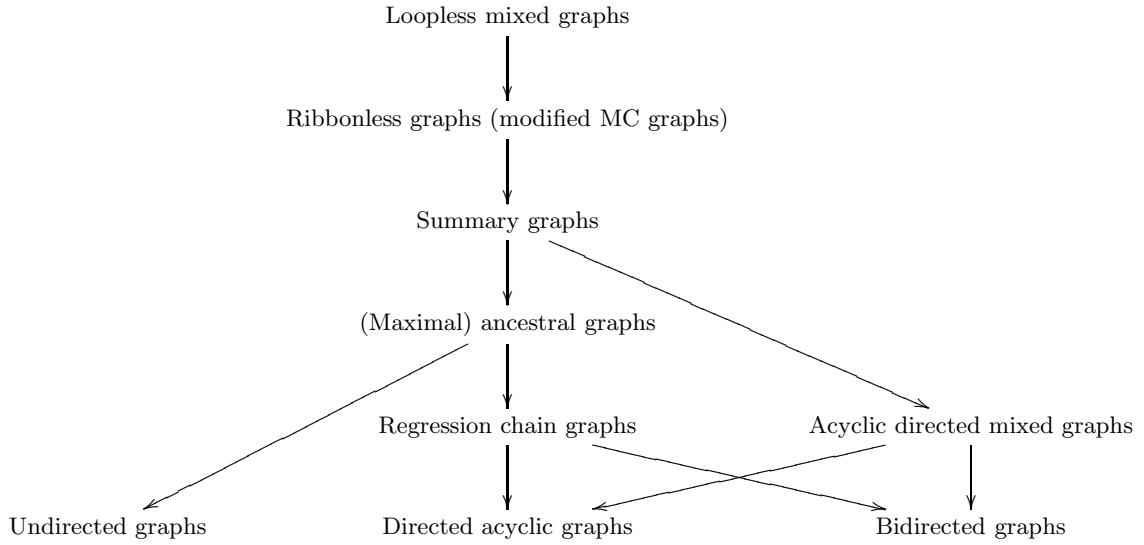


Figure 1: The hierarchy of subclasses of LMGs.

The common feature of all these subclasses is that these use the same interpretation for independencies, known as the m -separation criterion. We will shortly introduce the m -separation criterion for MAGs and its subclasses.

Ribbonless graphs, defined and studied in Sadeghi (2011) is a modification of MC graphs, defined by Koster (2002), to discard the line loops and to use the m -separation criterion.

Definition of ancestral graphs and regression chain graphs. An *ancestral graph* is a mixed graph that, for all nodes i , has (1) $i \notin \text{an}(\text{pa}(i) \cup \text{sp}(i))$ and (2) If $\text{ne}(i) \neq \emptyset$ then $\text{pa}(i) \cup \text{sp}(i) = \emptyset$. This means that there is no arrowhead pointing to a line and there is no direction-preserving cycle, and there is no arc with one endpoint the ancestor of the other endpoint, in the graph.

A graph $G = (V, E)$ is a *regression chain graph* if it contains at most the three types of edge, there is no arrowhead pointing to a line in graph, and it does not contain any *arc-direction-preserving* cycle, i.e. a cycle that contains arcs and at least one arrow and whose arrows are all towards one direction. Thus in such graphs the subgraph induced by lines is so-called at the beginning of graph, and the subgraph induced by the arrows and arcs is characterised by having a node set that can be partitioned into numbered subsets forming so-called *chains*, i.e. $V = \tau_1 \cup \dots \cup \tau_T$ such that all edges between nodes in the same subset are arcs and all edges between different subsets are arrows pointing from the set with the higher number to the one with the lower number. One can observe that in the subgraph induced by the arrows and arcs if we replace every τ_i by a node, we get a DAG.

The m -separation criterion Since, as we shall see, we are only interested in the subclasses of MAGs, we use the simplified version of m -separation criterion, defined in Sadeghi & Lauritzen (2011). This is identical to the original definition of m -separation; see Richardson & Spirtes (2002).

Let C be a subset of the node set V of a MAG. A path is m -connecting given C if all its collider nodes are in $C \cup \text{an}(C)$ and all its non-collider nodes are outside C . For two other disjoint subsets of the node set A and B , we say $A \perp_m B \mid C$ if there is no m -connecting path between A and B given C .

Notice that the m -separation criterion gives an interpretation of independencies on graphs, i.e. it induces an independence model.

4 Basic concepts for Markov equivalence

Definition of Markov equivalence. Now we can formally define Markov equivalence. Two graphs $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ are *Markov equivalent* if, for all subsets A , B , and C of V , $A \perp_m B \mid C$ in G_1 if and only if $A \perp_m B \mid C$ in G_2 .

Maximality and Markov equivalence. A loopless mixed graph G is called *maximal* if by adding any edge to G the independence model induced by the m -separation criterion changes. Alternatively, a graph $G = (V, E)$ is maximal if and only if, for every pair of non-adjacent nodes i and j of V , there exists a subset C of $V \setminus \{i, j\}$ such that $i \perp_m j \mid C$; see Richardson & Spirtes (2002); Sadeghi & Lauritzen (2011).

This implies that two Markov equivalent maximal graphs must have the same skeleton, where the *skeleton* of a graph results by replacing each edge by a full line.

Motivations behind using MAGs and its subclasses. In this paper we aim to discuss Markov equivalence for the subclasses presented in Fig. 1. The conditions for internal Markov equivalence for RGs and SGs are very complex. However, in Sadeghi (2011) it was demonstrated how RGs can be mapped into a Markov equivalent SG, and how SGs can be mapped into a Markov equivalent AG. Notice that ADMGs are SGs without full lines, so by the same method one can map ADMGs into Markov equivalent AGs.

In addition, since Markov equivalent maximal graphs must have the same skeleton, conditions for Markov equivalence for MAGs are simpler than those for Markov equivalence for AGs. In Richardson & Spirtes (2002) it was shown how AGs can be mapped into a Markov equivalent MAG. Therefore, we map all types of stable independence graphs into MAGs and discuss the Markov equivalences for MAGs and its subclasses.

Notice that all subclasses of MAGs discussed here are maximal by nature. Therefore, for their Markov equivalence they must have the same skeleton.

5 Internal and external Markov equivalences

5.1 Internal Markov equivalence for maximal ancestral graphs

Thus far, there are two elegant results regarding Markov equivalence for MAGs available (Ali *et al.*, 2005; Zhao *et al.*, 2004). These results use different definitions (colliders with order and minimal collider paths) and arguments. Even though it is not immediately obvious from their formulations, it can be shown that these are equivalent.

First result for Markov equivalence for MAGs. In order to present the first theorem, we quote two definitions from Ali *et al.* (2005). A path $\pi = \langle j, q_1, q_2, \dots, q_m, l, i \rangle$, with j not adjacent to i , is a *discriminating path* for $\langle q_m, l, i \rangle$ in G if and only if, for every node q_n , $1 \leq n \leq m$ on π , i.e. excluding j , i , and l ,

- i) q_n is a collider on π ; and
- ii) $q_n \longrightarrow i$, hence forming a non-collider along the path $\langle j, q_1, \dots, q_n, i \rangle$.

Fig. 2 illustrates what a discriminating path looks like.

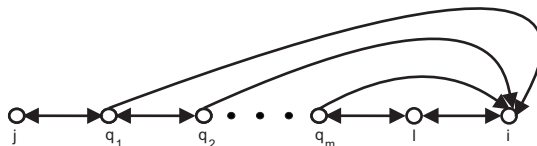


Figure 2: A discriminating path.

Let \mathcal{D}_n be the set of *triples of order n* defined recursively as follows:

Order 0: A triple $\langle h, l, i \rangle \in \mathcal{D}_0$ if h and i are not adjacent in G .

Order $n + 1$: A triple $\langle h, l, i \rangle \in \mathcal{D}_{n+1}$ if

- 1) $\langle h, l, i \rangle \notin \mathcal{D}_p$, for some $p < n + 1$ and
- 2) there exists a discriminating path $\pi = \langle j, q_1, \dots, q_m = h, l, i \rangle$ for l in G , and each of the colliders on the path: $\langle j, q_1, q_2, \rangle, \dots, \langle q_{m-1}, q_m, l \rangle \in \bigcup_{p \leq n} \mathcal{D}_p$.

If $\langle h, l, i \rangle \in \mathcal{D}_n$ then the triple is said to have order n . A discriminating path is said to have order n if every triple on the path has order at most n and at least one triple has order n . If a triple has order n for some n we then say that the triple has order, likewise for discriminating paths.

Theorem 1. (Ali & Richardson, 2004) *MAGs H_1 and H_2 are Markov equivalent if and only if H_1 and H_2 have the same skeleton and colliders with order.*

In Fig. 3 there are three MAGs with the same skeleton. In H_1 and H_2 since $i \sim k$, the collider $\langle i, j, k \rangle$ is with order 0, whereas in H_3 this is not a collider. In H_1 and H_2 , the collider $\langle j, k, h \rangle$ is with order 1. Therefore, we conclude that H_1 and H_2 are Markov equivalent, but these are not Markov equivalent to H_3 .

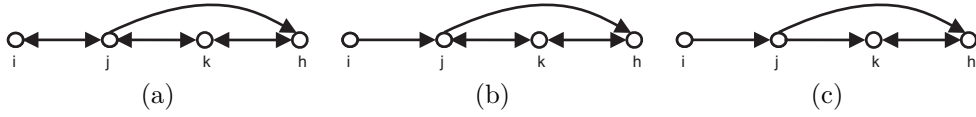


Figure 3: (a) A maximal ancestral graph H_1 . (b) A maximal ancestral graph H_2 that is Markov equivalent to H_1 . (c) A maximal ancestral graph H_3 that is not Markov equivalent to H_1 or H_2 .

Second result for Markov equivalence for MAGs. In order to present the second theorem, we quote two definitions from Zhao *et al.* (2004). A path π is called a *collider path* if all its inner nodes are colliders on π . A collider path $\pi = \langle i, B, j \rangle$ is called a *minimal collider path* if $i \not\sim j$ and there is no proper subset $B' \subset B$ such that $\langle i, B', j \rangle$ is a collider path between i and j . If $i \sim j$ then we call π a *minimal collider cycle*. In the graph in Fig. 4 the path $\langle i, j, k, h \rangle$ is a collider path, but it is not minimal collider since there exists the collider path $\langle i, j, h \rangle$, which is minimal collider.

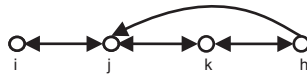


Figure 4: A non-minimal collider path $\langle i, j, k, h \rangle$ and a minimal collider path $\langle i, j, h \rangle$.

Theorem 2. (Zhao et al., 2004) *MAGs H_1 and H_2 are Markov equivalent if and only if H_1 and H_2 have the same skeleton and minimal collider paths.*

For the graphs in Fig. 3, by Theorem 2 we can make the same conclusion as before. To observe this it is enough to check that $\langle i, j, k \rangle$ and $\langle i, j, k, h \rangle$ are the minimal collider paths of H_1 and H_2 , whereas there is no minimal collider path in H_3 . We, therefore, conclude that H_1 and H_2 are Markov equivalent, but these are not Markov equivalent to H_3 .

5.2 Internal Markov equivalence for subclasses of maximal ancestral graphs

Markov equivalence for DAGs. First of all we recall a well-known result regarding Markov equivalence for DAGs.

Proposition 1. (Verma & Pearl, 1990; Frydenberg, 1990) *DAGs G_1 and G_2 are Markov equivalent if and only if they have the same skeleton and unshielded collision V-configurations.*

In the example in Fig. 5 all three graphs have the same skeleton. In G_1 there are two unshielded collision V-configurations $\langle k, i, h \rangle$ and $\langle j, i, h \rangle$. In G_2 there are the same unshielded collision V-configurations. Therefore, these two graphs are Markov equivalent. The only unshielded collision V-configuration in G_3 is, however, $\langle k, i, h \rangle$. Hence this graph is not Markov equivalent to G_1 and G_2 .

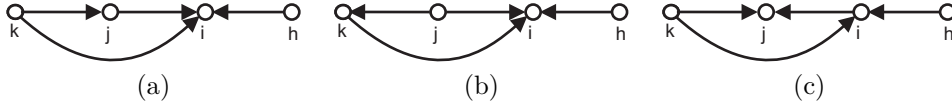


Figure 5: (a) A directed acyclic graph G_1 . (b) A Markov equivalent directed acyclic graph G_2 to G_1 . (c) A directed acyclic graph G_3 that is not Markov equivalent to G_1 .

Markov equivalence for UGs and BGs. The following proposition shows when two bidirected or undirected graphs are Markov equivalent.

Proposition 2. *Bidirected or undirected graphs H_1 and H_2 are Markov equivalent if and only if they are equal.*

Proof. In the case of the undirected graph, the result follows from Theorem 1 and the fact that there is no collider in undirected graphs.

In the case of the bidirected graph, the result follows from Theorem 2 and the fact that every path in bidirected graphs is a collider path. \square

Markov equivalence for RCGs. Since RCGs are a subclass of MAGs, we simplify the conditions for Markov equivalence for MAGs in order to obtain the conditions for Markov equivalence for RCGs.

Proposition 3. (*Wermuth & Sadeghi, 2011*) *RCGs H_1 and H_2 are Markov equivalent if and only if H_1 and H_2 have the same skeleton and unshielded collider V-configurations.*

Proof. We apply Theorem 1 to RCGs and simplify its conditions in order to obtain the conditions of this theorem. The first condition of Theorem 1 (having the same skeleton) is the same as the first condition of this theorem. Therefore, it is enough to prove that H_1 and H_2 have the same colliders with order if and only if they have the same unshielded collider V-configurations.

An unshielded collider V-configuration is by definition a collider with order. We prove that in RCGs a collider V-configuration that is a collider with order is unshielded. This proves the proposition: Suppose that $\langle h, k, l \rangle$ is a collider with order and, for contradiction, is not unshielded. By the definition of collider with order there exists a discriminating path $\langle x, q_1, \dots, q_p = h, k, l \rangle$ for k . Hence $h \in \text{sp}(k)$. In addition, if $l \in \text{pa}(k)$ then $h \in \text{an}(k)$, a contradiction by the definition of RCGs. Therefore, $l \in \text{sp}(k)$, but again this is a contradiction since in RCGs for a collider V-configuration with two adjacent arcs $\langle h, k, l \rangle$, one endpoint (h) cannot be the parent of the other endpoint (l). \square

In the example in Fig. 6 all three RCGs have the same skeleton. In H_1 there are three unshielded collider V-configurations $\langle l, h, k \rangle$, $\langle l, j, i \rangle$, and $\langle k, i, j \rangle$. In H_2 there are the same unshielded collider V-configurations. Therefore, these two graphs are Markov equivalent. The unshielded collider V-configurations in H_3 are, however, $\langle l, h, k \rangle$ and $\langle k, i, j \rangle$. Hence this graph is not Markov equivalent to H_1 or H_2 .

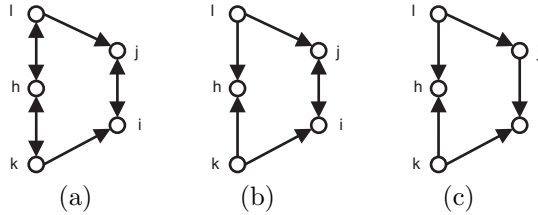


Figure 6: (a) A regression chain graph H_1 . (b) A Markov equivalent regression chain graph H_2 to H_1 . (c) An RCG (DAG) H_3 that is not Markov equivalent to H_1 .

5.3 External Markov equivalence for subclasses of maximal ancestral graphs

As a corollary of Proposition 3, in order to check the external Markov equivalence for every two introduced subclasses of MAGs (excluding MAGs), i.e. RCGs, BGs, UGs,

and DAGs, the conditions that are used for Markov equivalence for RCGs can be used. In some cases, if we suppose that the graphs satisfy the conditions for representational Markov equivalence, which will be introduced in the next section, then the conditions for external Markov equivalence can be simplified.

Corollary 1. *Every two of RCG, BG, UG, and DAG are Markov equivalent if and only if they have the same skeleton and unshielded collider V-configurations.*

Proof. The result follows from the fact that BGs, UGs, and DAGs are subclasses of RCGs. \square

Notice that for Markov equivalence for a UG and a graph H of another type the corollary states that there should be no collider V-configurations in H .

6 Representational Markov equivalence and algorithms

Structure of the section. In this section, in each subsection, we deal with representational Markov equivalence for MAGs and its subclasses to a specific subclass. In each subsection we first introduce an algorithm for generating a graph of a specific subclass which is Markov equivalent to a given MAG. The algorithm is usually trivially simplified for subclasses of MAGs. We then introduce conditions for a MAG under which it is Markov equivalent to a graph from the given subclass. By simplifying these conditions we obtain the conditions for subclasses of MAGs under which they are Markov equivalent to a graph of the given subclass. Notice that representational Markov equivalence to the class of MAGs produces trivial results.

6.1 Representational Markov equivalence to directed acyclic graphs

Generating a DAG which is Markov equivalent to a given MAG. We begin with an algorithm for generating a Markov equivalent DAG to a given MAG that satisfies the conditions of Lemma 1.

Algorithm 1. *(Generating a Markov equivalent DAG to a maximal ancestral graph H)*
Start from H .

1. Apply the maximum cardinality search algorithm on $H[-]$ to order the nodes.
2. Orient the edges of $H[-]$ from a lower number to a higher one.
3. Replace unshielded collider V-configurations by unshielded collision V-configurations, i.e. replace $i \leftrightarrow o \leftrightarrow j$ and $i \leftrightarrow o \leftarrow j$ by $i \rightarrow o \leftarrow j$ when $i \not\sim j$.
4. Order the nodes of the subgraph induced by arrows such that arrows are from higher numbers to lower ones.

5. Order the nodes of the subgraph induced by arcs arbitrarily by numbers not used in the previous step if the number for the node does not already exist.

6. Replace arcs by arrows from higher numbers to lower ones.

Continually apply each step until it is not possible to apply the given step further before moving to the next step.

Lemma 1. *Let H be a maximal ancestral graph. If $H[—]$ is chordal and there is no minimal collider path or cycle of length 4 in H then Algorithm 1 generates a Markov equivalent DAG to H .*

Proof. Denote the generated graph by G . Graph G is directed since by Algorithm 1, all edges are turned into arrows. Since $H[—]$ is chordal, the graph generated by the perfect elimination order of the maximal cardinality search does not have a direction-preserving cycle; see Section 2.4, Blair & Peyton (1992).

In addition, the arrows present in the graph do not change by the algorithm. We show that there is no direction-preserving cycle after applying step 3: If, for contradiction, there is a shortest direction-preserving cycle after applying step 3 then a collider V-configuration $\langle j, k, i \rangle$ (say the jk -edge is an arc) should turn into a transition V-configuration after applying step 3. In this case there is an hj -edge for a node h with an arrowhead pointing to j and $h \not\prec k$. Since there is no minimal collider path or cycle of length 4, $\langle j, k, i \rangle$ is shielded. Notice that on the ji -edge there is an arrowhead pointing to j since otherwise there is a minimal collider path or cycle of length 4. This implies that the ji -edge is an arc, since otherwise a shorter direction-preserving path via the arrow from i to j is generated.

Since this edge should turn into an arrow from j to i , $i \sim h$ and there is a node l pointing to i such that $l \not\prec j$. Since $\langle h, j, i, l \rangle$ and $\langle k, j, i, l \rangle$ are collider paths (or cycles), on the hi -edge there is an arrowhead pointing to i , and ki is an arc. To turn the ki -arc into an arrow from i to k there is a kp -edge with its arrowhead pointing to k , and $p \not\prec i$. Therefore, $\langle p, k, i, h \rangle$ is a minimal collider path (or cycle), a contradiction. Therefore, there is no direction-preserving cycle after applying step 3.

Therefore, the ordering of step 4 is permissible, and by step 6 there are obviously no direction-preserving cycles generated. Therefore, G is acyclic.

Now we prove that G is Markov equivalent to H : Since there is no minimal collider path of length 4 in H , by Theorem 2, H is Markov equivalent to G if and only if they have the same skeleton and unshielded collider V-configurations. Graph G obviously has the same skeleton as that of H . In addition, an unshielded collider V-configuration in G is an unshielded collider V-configuration in H . If, for contradiction, an unshielded collider V-configuration $\langle i, k, j \rangle$ in H does not exist in G then one of the arrowheads pointing to

k , say on edge ik , must be removed by step 3. Therefore, there is an unshielded collider V -configuration $\langle l, i, k \rangle$ in H . Now $\langle l, i, k, j \rangle$ is a minimal collider path since $l \not\sim k$ and $i \not\sim j$, a contradiction. \square

It is easy to see that, for UGs, BGs, and RCGs that can be Markov equivalent to DAGs, the algorithm generates a Markov equivalent DAG to a given RCG, steps 1 and 2 of the algorithm generate a Markov equivalent DAG to a given UG, and steps 3-6 of the algorithm generate a Markov equivalent DAG to a given BG.

Fig. 7 illustrates how to apply Algorithm 1 step by step to a MAG.

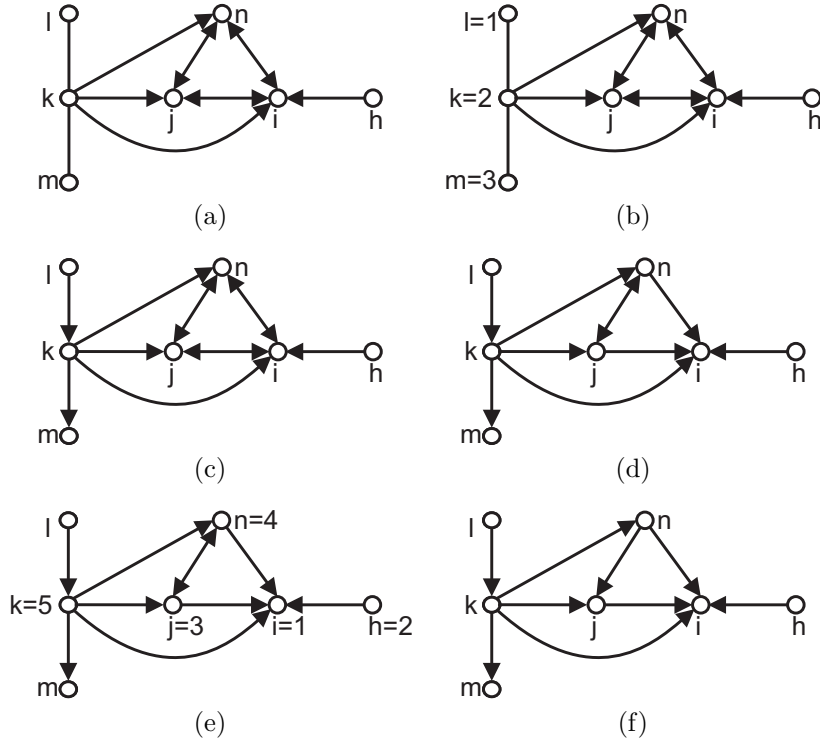


Figure 7: (a) A MAG. (b) The generated graph after applying step 1 of Algorithm 1. (c) The generated graph after applying step 2. (d) The generated graph after applying step 3. (e) The generated graph after applying step 4. (f) The generated DAG after applying step 6.

Conditions for representational Markov equivalence for the class of MAGs and its subclasses to DAGs. The following proposition shows that sufficient conditions for a given MAG, presented in Lemma 1, are also necessary. The corollaries of this proposition illustrate the conditions under which RCGs, BGs, and UGs can be Markov equivalent to a DAG.

Theorem 3. *A maximal ancestral graph H is Markov equivalent to a DAG if and only if $H[-]$ is chordal and there is no minimal collider path or cycle of length 4 in H .*

Proof. (\Rightarrow) Suppose that the maximal ancestral graph H is Markov equivalent to a directed acyclic graph G . By Theorem 2 G must have the same skeleton and minimal collider paths as those of H . Since there is no collider V-configuration in $H[-]$, the corresponding induced subgraph of G should have no unshielded collision nodes. If, for contradiction, this subgraph contains an induced C_n , $n \geq 4$, then there exists a collision V-configuration on the cycle, otherwise there exists a direction-preserving cycle in G , which is not permissible. This collision V-configuration is unshielded since $n \geq 4$ and C_n is chordless. This is a contradiction since G and H have the same skeleton. If H contains a minimal collider path or cycle π then edges of π cannot be oriented in G to generate a collider path.

(\Leftarrow) The result follows from Lemma 1. \square

We recall the following known statement for BGs as a corollary to the proposition.

Corollary 2. (*Pearl & Wermuth, 1994*) *A BG is Markov equivalent to a DAG if and only if it does not contain any P_4 or C_4 as induced subgraphs.*

Proof. For BGs, every path is a collider path, and every minimal collider path or cycle is a P_4 or C_4 . Using these, the result follows. \square

We also recall the following known statement for UGs as a corollary to the proposition; see Lauritzen (1996).

Corollary 3. *A UG is Markov equivalent to a DAG if and only if it is chordal.*

Proof. For UGs, $H[-] = H$, and there is no collider path in UGs. Using these, the result follows. \square

The following corollary shows the conditions under which RCGs can be Markov equivalent to DAGs.

Corollary 4. *An RCG with chain component node sets τ_1, \dots, τ_T is Markov equivalent to a DAG if and only if, $H[-]$ is chordal and for $1 \leq i \leq T$, the induced subgraph by $\tau_i \cup \text{pa}(\tau_i)$ does not contain any collider P_4 or C_4 as an induced subgraph.*

Proof. For RCGs, every collider path is in one of $\tau_i \cup \text{pa}(\tau_i)$, $1 \leq i \leq T$. In addition, in RCGs a minimal collider path or cycle is chordless. Using these, the result follows. \square

Necessary conditions for representational Markov equivalence to DAGs. Here we introduce some necessary conditions for Markov equivalence for MAGs, BGs, and RCGs to a DAG. For this purpose we use the following well-known graph theoretical result:

Lemma 2. *If a graph G contains no P_4 or C_4 as an induced subgraph then there is a node that is adjacent to all other nodes.*

Corollary 5. *Let H be a MAG and $H[\leftrightarrow] = (V, E)$. If H is Markov equivalent to a DAG then there exists a node that is adjacent to all other nodes in $V \cup \text{pa}(V)$.*

Proof. Graphs with no minimal collider paths or cycles of length 4 do not contain a collider P_4 or C_4 . In addition, every collider path in MAGs is in $V \cup \text{pa}(V)$. Using these, the result follows from Theorem 3 and Lemma 2. \square

Corollary 6. *If a bidirected graph is Markov equivalent to a DAG then there exists a node that is adjacent to all other nodes.*

Proof. The result follows from Corollary 2 and Lemma 2. \square

Corollary 7. *If an RCG is Markov equivalent to a DAG then in each $\tau_i \cup \text{pa}(\tau_i)$, $1 \leq i \leq T$, there exists a node that is adjacent to all other nodes in $\tau_i \cup \text{pa}(\tau_i)$.*

Proof. The result follows from Corollary 4 and Lemma 2. \square

6.2 Representational Markov equivalence to undirected and bidirected graphs

Generating a UG which is Markov equivalent to a given MAG. By removing all arrowheads one generates a Markov equivalent UG to a given MAG that satisfies the condition of Lemma 3.

Lemma 3. *For a maximal ancestral graph H with no unshielded collider V -configuration, removing arrowheads generates a Markov equivalent UG to H .*

Proof. The generated graph is obviously a UG and is also the only UG that has the same skeleton as that of H . Neither H nor the generated graph contains any minimal collider paths. This completes the proof. \square

One can therefore observe that, for DAGs, BGs, and RCGs that can be Markov equivalent to UGs, removing arrowheads generates a Markov equivalent UG to the given graph.

Conditions for representational Markov equivalence for MAGs and its subclasses to UGs. The following proposition shows that the sufficient condition for a given MAG, presented in Lemma 3, is also necessary. The corollaries of this proposition illustrate the conditions under which BGs and DAGs can be Markov equivalent to a UG.

Proposition 4. *A maximal ancestral graph H is Markov equivalent to a UG if and only if there is no unshielded collider V -configuration in H .*

Proof. (\Rightarrow) Suppose that H is Markov equivalent to an undirected graph G . Graphs H and G have the same skeleton and minimal collider paths, but G has no minimal collider paths. Since an unshielded collider V-configuration is a minimal collider path, H contains no unshielded collider V-configurations.

(\Leftarrow) The result follows from Lemma 3. \square

One can also use this result for RCGs. Here we simplify the condition further for DAGs and BGs.

Corollary 8. *A directed acyclic graph G is Markov equivalent to a UG if and only if there is no unshielded collision V-configuration in H .*

Proof. The result follows from the fact that the only type of colliders in DAGs is collisions. \square

Corollary 9. *A BG is Markov equivalent to a UG if and only if it is complete.*

Proof. The result follows from the fact that all unshielded V-configurations in BGs are collider. \square

Generating a BG which is Markov equivalent to a given MAG. Replacing all edges by arcs generates a Markov equivalent BG to a given MAG that satisfies the condition of Lemma 4.

Lemma 4. *For a maximal ancestral graph H with no unshielded non-collider V-configuration, replacing all edges by arcs generates a Markov equivalent BG to H .*

Proof. The generated graph is obviously a BG and is also the only BG that has the same skeleton as that of H . All V-configurations in both H and the generated graph are colliders. This completes the proof. \square

One can therefore observe that, for DAGs, UGs, and RCGs that can be Markov equivalent to BGs, replacing all edges by arcs generates a Markov equivalent BG to the given graph.

Conditions for representational Markov equivalence for MAGs and its subclasses to BGs. The following proposition shows that the sufficient condition for a given MAG, presented in Lemma 4, is also necessary.

Proposition 5. *A maximal ancestral graph H is Markov equivalent to a BG if and only if there is no unshielded non-collider V-configuration in H .*

Proof. (\Rightarrow) Suppose that H is Markov equivalent to a bidirected graph G . Graphs H and G have the same skeleton and minimal collider paths, but every minimal path in G is a minimal collider path. Since an unshielded non-collider V-configuration is a minimal but not a collider path, H contains no unshielded non-collider V-configurations.

(\Leftarrow) The result follows from Lemma 4. \square

One can also use this result for RCGs. Here we simplify the condition further for DAGs and UGs. The results have been known in the literature for long time, e.g., see Pearl & Wermuth (1994).

Corollary 10. *A directed acyclic graph G is Markov equivalent to a BG if and only if there is no unshielded non-collision V-configuration in H .*

Proof. The result follows from the fact that the only type of colliders in DAGs is collisions. \square

Corollary 11. *A UG is Markov equivalent to a BG if and only if it is complete.*

Proof. The result follows from the fact that all unshielded V-configurations in UGs are non-collider. \square

6.3 Representational Markov equivalence to regression chain graphs

Generating an RCG which is Markov equivalent to a given MAG. We begin with an algorithm for generating a Markov equivalent RCG to a given MAG that satisfies the conditions of Lemma 5.

Algorithm 2. *(Generating a Markov equivalent RCG to a MAG H)*

Start from H .

1. *For a non-collider V-configuration $i \leftarrow j \rightarrow k$ on an arc-direction-preserving cycle, remove the arrowhead pointing to j on the ij -edge when there is no unshielded collider V-configuration of form $\langle i, j, l \rangle$.*

Continually apply this step until it is not possible to apply it further.

Lemma 5. *For a maximal ancestral graph H with no arc-direction-preserving cycle on which every non-collider V-configuration $i \leftarrow j \rightarrow k$ is such that there is an unshielded collider V-configuration of form $\langle i, j, l \rangle$, Algorithm 2 generates a Markov equivalent RCG to H .*

Proof. Denote the generated graph by G . To show G is an RCG, it is enough to show that there is no arc-direction-preserving cycle in G . We know that the only arc-direction-preserving cycles in H are those on which there is a non-collider V-configuration

$i \leftarrow j \rightarrow k$ such that there is no unshielded collider V-configuration $\langle i, j, l \rangle$. In this case step 3 of the algorithm generates a source V-configuration, and therefore, destroys the arc-direction-preserving cycle.

We now prove that G is Markov equivalent to H : First, we prove that minimal collider paths in H remain unchanged in G . Suppose, for contradiction, that there is a minimal collider path π of length n , $n \geq 3$, containing an ij -arc, and the arrowhead pointing to j is removed by step 3 of the algorithm because of a V-configuration $i \leftarrow j \rightarrow k$ on an arc-direction-preserving cycle. Denote the three consecutive nodes on π by $\langle i, j, l \rangle$. Since there is an arrowhead pointing to j on the jl -edge, $i \sim l$. Since π is minimal collider, there exists another node on π , say h adjacent to l , i is an endpoint of π , and the li -edge is an arrow from l to i .

Now there is an arc-direction-preserving cycle $\langle i, j, l \rangle$ on which the only non-collider V-configuration $j \leftarrow l \rightarrow i$ is such that there is a collider V-configuration $\langle j, l, h \rangle$. Therefore, this collider V-configuration should be shielded, i.e. $h \sim j$. This edge is an arrow from j to h because π is minimal collider. Therefore, $\langle i, j, l, h \rangle$ is a primitive inducing path, and since H is maximal, $i \sim h$. This contradicts the fact that π is a minimal collider path. Therefore, minimal collider paths in H do not change by the algorithm.

In addition, a non-minimal collider path in H cannot turn into a minimal collider path in G , since we know that in RCGs all collider paths are chordless. This completes the proof. \square

Fig. 8 illustrates how to apply Algorithm 2 to a MAG.

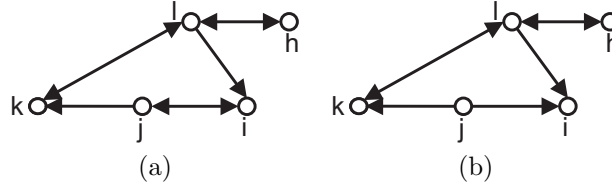


Figure 8: (a) A MAG. (b) The generated RCG after applying Algorithm 2.

Conditions for representational Markov equivalence for MAGs and its subclasses to RCGs. Since UGs, BGs, and DAGs are subclasses of RCGs, there are no conditions needed under which they are able to be Markov equivalent to an RCG. The following proposition shows when a given MAG can be Markov equivalent to an RCG. The corollary of this proposition shows when UGs can be Markov equivalent to an RCG.

Theorem 4. *A maximal ancestral graph H is Markov equivalent to an RCG if and only if there is no arc-direction-preserving cycle on which every non-collider V-configuration $i \leftarrow j \rightarrow k$ is such that there is an unshielded collider V-configuration of form $\langle i, j, l \rangle$.*

Proof. (\Rightarrow) Suppose that H is Markov equivalent to a multivariate regression chain graph G . Suppose, for contradiction, that there is an arc-direction-preserving cycle π' on which every non-collider V-configuration $i \leftarrow j \rightarrow k$ is such that there is an unshielded collider V-configuration $\langle i, j, l \rangle$. If π' has a chord qr then there are two shorter cycles including the chord and nodes on π' that are on one side of q and r . One can observe that at least one of the two cycles has the same property as the property of π' depending on whether there is an arrowhead pointing to q or r on ij -path. Hence, consider the minimal cycle π in this sense, which is chordless.

On π all collider V-configurations are unshielded. In addition, all collider V-configurations of form $\langle i, j, l \rangle$ are also unshielded. Therefore, since H is Markov equivalent to G , in G all these collider V-configurations should be preserved. Hence, all arcs on π remain arcs in G . Moreover, by replacing the arrows on π by arcs or by changing their directions a new unshielded collider V-configuration is generated. Therefore, arrows on π are also unchanged in G . Therefore, π exists in G . Since we know that arc-direction-preserving cycles are not permissible in RCGs, this is a contradiction.

(\Leftarrow) The result follows from Lemma 5. □

7 Summary

Summary of internal and external Markov equivalence for MAGs and its subclasses. In Section 5, we showed that for internal and external Markov equivalences for subclasses of MAGs, excluding MAGs themselves, the conditions for Markov equivalences for DAGs can be generalised naturally by using colliders instead of collisions. In other words, two subclasses of MAGs are Markov equivalent if and only if they have the same skeleton and unshielded collider V-configurations.

Summary of representational Markov equivalence for MAGs and its subclasses. The following table represents the summary of the conditions needed for representational Markov equivalence for MAGs and its subclasses. In addition, for each non-trivial case of table, we provided algorithms to generate a graph other types that is Markov equivalent to the graph of a given type.

The conditions presented in the table are for the graphs of the type indicated on the left column, which are to be Markov equivalent to a graph of the type indicated on the first row.

Table 1: Necessary and sufficient conditions on H a graph of a subclass of maximal ancestral graph on the left column to be able to be Markov equivalent to a graph of the subclass of maximal ancestral graph on the top row.

$H \backslash$	RCG	BG	UG	DAG
MAG	No arc-dir-pr cycle with every $i \leftrightarrow j \rightarrow k$ s.t. there is an unshielded collider $\langle i, j, l \rangle$	No unshielded non-collider.	No unshielded collider	$H[—]$ Chordal; no minimal collider path or cycle of length 4
RCG	-	No unshielded non-collider	No unshielded collider	No collider P_4 or C_4 in $\tau_i \cup \text{pa}(\tau_i)$
BG	-	-	Complete	No P_4 or C_4
UG	-	Complete	-	Chordal
DAG	-	No unshielded non-collision	No unshielded collision	-

Acknowledgments

The author is grateful to Steffen Lauritzen and Nanny Wermuth for helpful comments.

References

- Ali, R. A. & Richardson, T. (2004). Searching across Markov equivalence classes of maximal ancestral graphs. In *Proceedings of the joint statistical meeting of the American statistical association*. Toronto, Canada.
- Ali, R. A., Richardson, T. & Spirtes, P. (2009). Markov equivalence for ancestral graphs. *Ann. Statist.* **37**, 2808–2837.
- Ali, R. A., Richardson, T., Spirtes, P. & Zhang, J. (2005). Towards characterizing Markov equivalence classes for directed acyclic graphs with latent variables. In *Proceedings of the proceedings of the twenty-first conference annual conference on uncertainty in artificial intelligence (uai-05)*. AUAI Press, Arlington, Virginia, pp. 10–17.
- Blair, J. R. S. & Peyton, B. (1992). An introduction to chordal graphs and clique trees. *Graph Theory and Sparse Matrix Computation (eds. George, A., Gilbert, J. R., Liu, J. W.)*, 1–29.
- Cox, D. R. & Wermuth, N. (1993). Linear dependencies represented by chain graphs (with discussion). *Stat. Sci.* **8**, 204–218; 247–277.
- Darroch, J. N., Lauritzen, S. L. & Speed, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *Ann. Statist.* **8**, 522–539.
- Frydenberg, M. (1990). The chain graph Markov property. *Scand. J. Stat.* **17**, 333–353.

- Kiiveri, H., Speed, T. P. & Carlin, J. B. (1984). Recursive causal models. *J. Aust. Math. Soc., Ser. A* **36**, 30–52.
- Koster, J. T. A. (2002). Marginalizing and conditioning in graphical models. *Bernoulli* **8**, 817–840.
- Lauritzen, S. L. (1996). *Graphical models*. Clarendon Press, Oxford, United Kingdom.
- Pearl, J. & Wermuth, N. (1994). When can association graphs admit a causal interpretation? *Models and Data, Artificial Intelligence and Statistics* **4**, 205–214.
- Richardson, T. S. & Spirtes, P. (2002). Ancestral graph Markov models. *Ann. Statist.* **30**, 962–1030.
- Sadeghi, K. (2011). Stable classes of graphs containing directed acyclic graphs. *submitted*.
- Sadeghi, K. & Lauritzen, S. L. (2011). Markov properties for loopless mixed graphs. *submitted*, <http://arxiv.org/abs/1103.2523>.
- Spirtes, P., Richardson, T. & Meek, C. (1997). The dimensionality of mixed ancestral graphs. Tech. Rep. CMU-PHIL-83, Philosophy Department, CMU.
- Spirtes, P. & Richardson, T. S. (1997). A polynomial-time algorithm for determining dag equivalence in the presence of latent variables and selection bias. In *Preliminary papers, sixth international workshop on ai and statistics*. Ft. Lauderdale, FL, USA, pp. 489–501.
- Tarjan, R. E. & Yannakakis, M. (1984). Simple linear-time algorithms to test chordality of graphs, test acyclicity of hypergraphs, and selectively reduce acyclic hypergraphs. *SIAM J. Comput.* **13**, 566–579.
- Verma, T. & Pearl, J. (1990). On the equivalence of causal models. In *Proceedings of the proceedings of the sixth conference annual conference on uncertainty in artificial intelligence (uai-90)*. Elsevier Science, New York, NY, pp. 220–227.
- Wermuth, N. (2011). Probability distributions with summary graph structure. *Bernoulli* **17**, 845–879.
- Wermuth, N. & Cox, D. R. (1998). On association models defined over independence graphs. *Bernoulli* **4**, 477–495.
- Wermuth, N. & Cox, D. R. (2004). Joint response graphs and separation induced by triangular systems. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **66**, 687–717.

Wermuth, N. & Sadeghi, K. (2011). Sequences of regressions and their independences. *TEST*, To appear as an invited discussion paper, <http://arxiv.org/abs/1103.2523>.

Zhao, H., Zheng, Z. & Liu, B. (2004). On the Markov equivalence of maximal ancestral graphs. *Sci. China Ser. A* **48**, 548–562.

Kayvan Sadeghi, Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG, UK.

Email: sadeghi@stats.ox.ac.uk